

FISHNet: encouraging data sharing and reuse in the freshwater science community

Abstract

This paper describes the FISHNet project, which developed a repository environment for the curation and sharing of data relating to freshwater science, a discipline whose research community is distributed thinly across a variety of institutions, and usually works in relative isolation as individual researchers or within small groups. As in other “small sciences”, these datasets tend to be small and “hand-crafted”, created to address particular research questions rather than with a view to reuse, so they are rarely curated effectively, and the potential for sharing and reusing them is limited. The paper addresses a variety of issues and concerns raised by freshwater researchers as regards data sharing, describes our approach to developing a repository environment that addresses these concerns, and identifies the potential impact within the research community of the system.

Keywords: freshwater; biology; data sharing; data publication; data reuse; data repositories; DOI; Fedora.

Introduction

The primary motivation for the FISHNet project was to facilitate the effective curation and sharing of data within the freshwater science community. The freshwater science community (at least in the UK, where the work described was based) is distributed thinly across a variety of academic institutions and departments, usually working in relative isolation as individual researchers or within small groups. Datasets created by these researchers tend to be relatively small and “hand-crafted”, created to address particular research questions. As a result, the datasets themselves are also distributed in small amounts across different institutions, which rarely have the facilities for curating them effectively. It is not easy to find or access such datasets, which affects the potential for reusing them for other research. As researchers typically create these datasets for their own use, there is little effort to apply good practice, or to take the potential for reuse, or even for longer-term curation, into account. Moreover, there are a number of cultural issues and concerns that militate against the sharing of data with other researchers. The broad aim of the FISHNet project was to implement and sustain a repository environment to facilitate sharing of and access to data within the freshwater science community, and to ensure the long-term curation and preservation of this data.

FISHNet was a collaboration between King’s College London (KCL) and the Freshwater Biological Association (FBA), a UK-based charity whose remit includes the support of freshwater research and the provision of data and information of interest to freshwater researchers.

The Community and its Data

The UK freshwater science community is thinly spread across different academic institutions, as for the most part freshwater scientists work within larger departments, such as biology, geography or health sciences, as a minority of the researchers. While some datasets are managed by non-academic organisations such as the UK Environment Agency, these tend to be larger datasets covering longer time periods, and the majority of freshwater datasets are relatively small and “hand-crafted”, created by the efforts of these scattered scientists in order to address particular issues. The majority of these datasets are gathered through field experiments, although they can also be

sourced from automated instrumentation such as scientific buoys stationed in a lake, or via measurements and analyses taken in the laboratory.

This has a number of consequences: for the most part, freshwater science data is widely distributed in small amounts across different institutions, and there is little in the way of policies or practices for curating this data – datasets are often managed in *ad hoc* ways on local servers or separate hard drives – or mechanisms for sharing the data. Moreover, a dataset is typically created by an individual freshwater scientist to address a particular research question, so there is little effort to make these datasets more reusable by others, e.g. by standardisation, and indeed resource limitations mean that such activities are considered to be of lower priority (see also Shearer (2009) and Key Perspectives Ltd (2010)). Nevertheless, the data collected would in many cases be great value for answering other research questions, especially if it could be combined with other datasets.

Objectives

Broadly speaking, the aim of the FISHNet project was to implement and sustain a repository environment to facilitate sharing of and access to data within the freshwater science community, and to ensure the long-term curation and preservation of this data. We planned to achieve this by the following specific objectives:

- Provide a community-centric environment for the discovery, sharing and reuse of data with other potential users in the research and academic communities.
- Provide services for data management, curation, discovery and sharing that are driven by the needs and practices of the freshwater research community.
- Build a sustainable infrastructure, backed up with a business plan, that leverages the cross-institutional support of universities and other research institutions, as well as the broad freshwater science community.
- Educate the freshwater science community as regards standards and good practice for creating and curating datasets, and provide an environment that helps the community to apply these standards and good practice.
- Facilitate the application and disseminations of discipline-specific standards that *emerge* in this community.
- Encourage sharing of datasets within the community, e.g. by offering incentives such as data curation guarantees and facilitating citation for datasets.
- Lowering the barrier to deposit by using a simple categorisation of access rights.

Established Practices and Challenges

Freshwater data is created or captured via a variety of methods depending on the research topic. Data may be reused from an existing data source, created via measurements taken in the lab – either entered directly into a digital format or added to a notebook and then transcribed – or collected from field experiments. Field experiments account for the majority of the data used by the researchers with whom FISHNet engaged, and, as with lab data, it may be entered directly into a computer, hand written and then transcribed, or sourced from automated instrumentation such as scientific buoys stationed in a lake.

The primary motive for creating FISHNet was to facilitate the sharing of data and to increase access to it for freshwater researchers. On the whole, there is a trend within the discipline for larger institutions such as the Centre for Ecology & Hydrology (CEH) and the Environment Agency (EA) to hold and maintain large datasets that represent readings taken over a long period, whereas in

smaller institutions individual freshwater scientists possess smaller datasets that focus on specific research topics. Some smaller institutions also make use of the data held by EA and CEH, which they tend to adapt and enhance with additions or corrections. However there exists no formal mechanism for unifying all these changes and additions and feeding them back into the datasets managed by EA and CEH. A good example of a dataset that is frequently used and modified is the EA's national ecology database, BIOSYS.

Freshwater scientists report that, in spite of data being in principle freely available from institutions such as the EA and CEH, in practice obtaining such data can be difficult, for several reasons. In many cases, it is simply not clear who in an organisation to contact about a particular dataset. There is also a lack of a clear institutional policy regarding data ownership and data IPR issues. Among the scientists with whom FISHNet engaged were individuals who stated that they refuse to share any data, and even when reminded that the data was obtained via public funding continued to maintain the position that the data belonged to them. In one particular case a Freedom of Information request had been submitted in order to oblige the researcher to release their data publically. This is not the case for all researchers, however, and there are many who are happy to release their data when asked to do so, but also some who only have the data in notebooks and do not have the time or resources to create digital versions of the data.

Data *curation* practice requires a clear understanding of *research* practice (Lyon et al., 2010; Borgman (2007)), so we engaged through interviews and questionnaires with a broad range of stakeholders in the field of freshwater science, both academic and non-academic (e.g. commercial, charitable and government). The user engagement exercise taught us that, within the freshwater science community, there were a variety of pressing concerns and issues regarding the sharing and the reuse of research data that needed to be addressed by a project such as FISHNet¹. These may be summarised as follows:

1. While adequate for the creator's own research purposes, the data may not have been prepared to a sufficiently high quality to be used by others. Typically, data collected for small projects or by individual researchers is recorded to address the needs of the individual project, or to answer a particular research question, without considering data reusability. Difficulties may occur in interpreting such data, as the dataset may contain implicit assumptions that are obvious to the original creator but may not be so to other users. The following list indicates the variety of ways in which the datasets examined failed to meet these criteria of reusability, although a broader examination might reveal others:
 - Non-standard column names.
 - Idiosyncratic abbreviations.
 - Units of measurement omitted or assumed.
 - Lack of metadata.
 - Use of non-standard metadata terms.
 - Lack of information about how the data was collected and for what purpose.
 - Existence of multiple, slightly different, copies of a dataset in circulation among researchers (i.e. no master copy and no versioning).
 - Inconsistent taxonomic resolution across datasets, e.g. a dataset may record information at the level of the taxonomic family, whereas recording to the species level is required to make it universally reusable.

¹ To put these concerns in a wider context, compare the discussions in Borgman et al. (2007) and Borgman (2010).

- Use of non-standard geographical terms or information, e.g. use of eastings and northings rather than latitude and longitude or Ordnance Survey grid reference².
2. If a dataset is made available too soon, then other researchers may pre-empt them by publishing work based on this data before the creator of the dataset can do so themselves.
 3. There is a lack of policy regarding data ownership and data IPR issues, which discourages researchers from sharing data.
 4. Lack of an effective mechanism for sharing data. Currently if data is shared it is on an individual basis, e.g. by emailing a dataset as an attachment. This has downsides, in particular arising from a lack of oversight once a dataset has been passed on – for example, the data may be forwarded to others, and it may be modified by someone else, leading to the existence of multiple versions. On the other hand, if there were a more effective system – such as that to be provided by the FISHNet project – then it should not disrupt the researchers work practices too much or it would not be used.
 5. In particular, if a dataset is shared, other people may use it in their research and thus gain academic credit – and perhaps a competitive advantage – without crediting the researcher that created the dataset. Ensuring that a researcher receives proper academic credit for their work in collecting and maintaining a dataset was a major issue, especially given the government framework for evaluating academic research in the UK, the Research Excellence Framework³, which is focused on researchers’ publication record and broader impacts
 6. Similarly, there was a concern that data might be used for commercial purposes without appropriate financial compensation to the dataset’s creator.
 7. Concern about having responsibility for maintaining the master copy of a dataset, if it is of particular importance and likely to be requested by other users on a frequent basis.
 8. While some researchers would be happy to share data freely, others wanted ultimate control over who accessed their data, and a few declined to share data at all, even when reminded that the data was obtained via public funding and there was an obligation to provide access to it.
 9. Concerns about privacy and confidentiality. These may occur, for example, in survey data that includes details about farms and land-owners, such as names and locations.

Addressing the Issues

Summary

Table 1 summarises how the FISHNet project addressed each of these issues in turn. The table refers to key components of the system – the “traffic light” system, issue of DOIs, and the simplified ingest workflow – that are described in detail in the subsequent subsections.

	<i>Issue</i>	<i>Addressed by</i>
1.	Data or metadata quality insufficient for reuse	Provision of expert assistance, ingest tools to help in the creation and validation of datasets, guidelines for best practice.
2.	Other researchers may use data to pre-empt research of data creator	Different levels of access can be selected. By choosing the Orange category, access can be restricted to researchers of the depositor’s choosing. It is possible to move between categories, so the

² Ordnance Survey is a UK government body that provides the national mapping agency for Great Britain.

³ <http://www.hefce.ac.uk/research/ref/>

		Orange category can be used to implement an embargo on using the data until the depositor is finished with it, at which point it can be moved to Green.
3.	Lack of IPR policy	The “traffic light” deposit system simplifies IPR, either allowing a single open licence, or else leaving access control to the data owner.
4.	Lack of effective sharing mechanism	FISHNet repository provides this mechanism. Simplified ingest process and traffic light system lowers barriers to deposit as much as possible.
5.	Loss of credit for data when data is reused by other researchers	The depositor can retain control over who accesses and uses the data by depositing in the Orange category. Provision of DOIs will facilitate citation of datasets and thus traceability – this not only helps with crediting their work, it is also an encouragement to deposit. By providing a single point of access, the repository access logs can be used to record downloads of the dataset.
6.	Lack of compensation when data reused for commercial purposes	Ditto.
7.	Responsibility for maintaining dataset	FISHNet repository takes responsibility for maintaining master dataset, including digital preservation activities where appropriate.
8.	Desire for ultimate control over access to dataset	Orange category allows user to retain control over access to dataset.
9.	Privacy and confidentiality	Anonymisation used in disseminated versions of datasets where necessary (original versions kept but access restricted).

Table 1: How FISHNet addressed users' concerns

“Traffic Light” system for deposit

The deposit of a dataset into the FISHNet repository falls into one of three categories, in which increasing levels of benefit are offered to the depositor in return for an increasing willingness to share the dataset. This is thus a “carrot” to encourage openness and data sharing. By using this approach we have also simplified the IPR situation, avoiding the need for a depositor to choose from a confusing variety of licences.

The three categories are labelled Red, Orange and Green, and thus we refer to this mechanism as the “traffic light” system. The implications of the three categories are as follows.

In the **Red** category, access is controlled by the data owner, who must provide basic metadata about the dataset, including contact details for people to request the dataset. If the dataset is already held in another repository, the data itself is not required – only a persistent URL for it. On the other hand, if a copy of the data *is* deposited with FISHNet, the data is not required to be in a readily reusable format, and the repository carries out no QA (Quality Assurance) and will carry out bit-level preservation only. Users can find the dataset in the catalogue but have to contact the data owner to obtain a copy, and rights are negotiated bilaterally between the owner and the requesting user.

In the **Orange** category, access is again controlled by the data owner. However, in this case the depositor must provide more detailed metadata, where the metadata schema may depend on the nature of the data, as well as making dataset available in a form deemed to be reusable by others.

To this end, FISHNet will provide expert assistance and, at a later date, tools (such as spreadsheet templates) for getting datasets into a suitable condition; this functionality will be built into the ingest workflow. Once the dataset is deemed to be in a reusable form a DOI will be assigned (see the following section on DOIs and Quality Assurance). FISHNet will also undertake to preserve the data in the long term. Just as for Red, users can find Orange datasets in the catalogue but have to contact the data owner to obtain a copy, and rights are negotiated bilaterally between the owner and the requesting user.

Datasets in the **Green** category must in addition be freely accessible from the repository, and the data owner has to make the dataset available under an appropriate licence. This is expected to be one of the Open Data Commons⁴ licences, as these were considered to be more appropriate for data than the Creative Commons⁵ licences, which are aimed rather at licensing creative works. In addition, Green datasets may be transformed into RDF and included in a triple store, so it can be exposed as linked data and semantically queried, and (in the longer term) tools may be integrated to allow owners to carry out this mapping for themselves.

DOIs and Quality Assurance

FISHNet addressed the issue of improving data reusability by carrying out quality assurance on data submissions that fall into the Orange or Green categories. Repository staff will work with the uploader of the dataset to ensure that it conforms to appropriate standards and is readily interpretable by other users. As this will in general require additional effort on the part of the uploader, some form of “reward” or “carrot” to motivate them will be required. FISHNet will achieve this by providing a DOI (see Brase, 2009) for the dataset, which will facilitate citation of the dataset and has the potential to increase significantly the number of citations they receive per publication and thus the visibility and ranking of the depositor’s research. This functionality is not yet implemented, although we have been in discussion with DataCite⁶ and we expect it to be added sometime during 2012.

One issue raised by the use of DOIs was the approach to take with datasets that are updated frequently, for example sensor datasets that may be updated many times per day. One aim of data publication using DOIs is to make datasets citable in publications. However, if a dataset is cited in a paper, there needs to be some assurance that the data contained in the dataset remains unchanged, both for peer review of the paper and for post-publication of published results by others. In any case, the rules for the use of DOIs require that the referenced object remains essentially unchanged for an indefinite period of time. Should a frequently updated dataset be assigned a new DOI every time it is updated? The approach taken was to think of a DOI as an ISBN for an edition of the dataset; this directly affected the design for the “splash page” of a dataset in FISHNet, as it implies that a single dataset can have multiple DOIs – corresponding to different editions – and that a DOI must resolve to a page that makes it clear which edition is being returned.

Ingest Workflow

These principles and components were built into an ingest workflow, implemented as part of the FISHNet user interface, to make the deposit process as painless as possible. Figure 1 represents the FISHNet ingest workflow using BPMN (Business Process Management Notation)⁷ notation. The repository itself was implemented using the Fedora Commons software⁸.

⁴ <http://opendatacommons.org/licenses/>

⁵ <http://creativecommons.org/licenses/>

⁶ <http://datacite.org/>

⁷ <http://www.bpmn.org/>

⁸ <http://www.fedora-commons.org/>

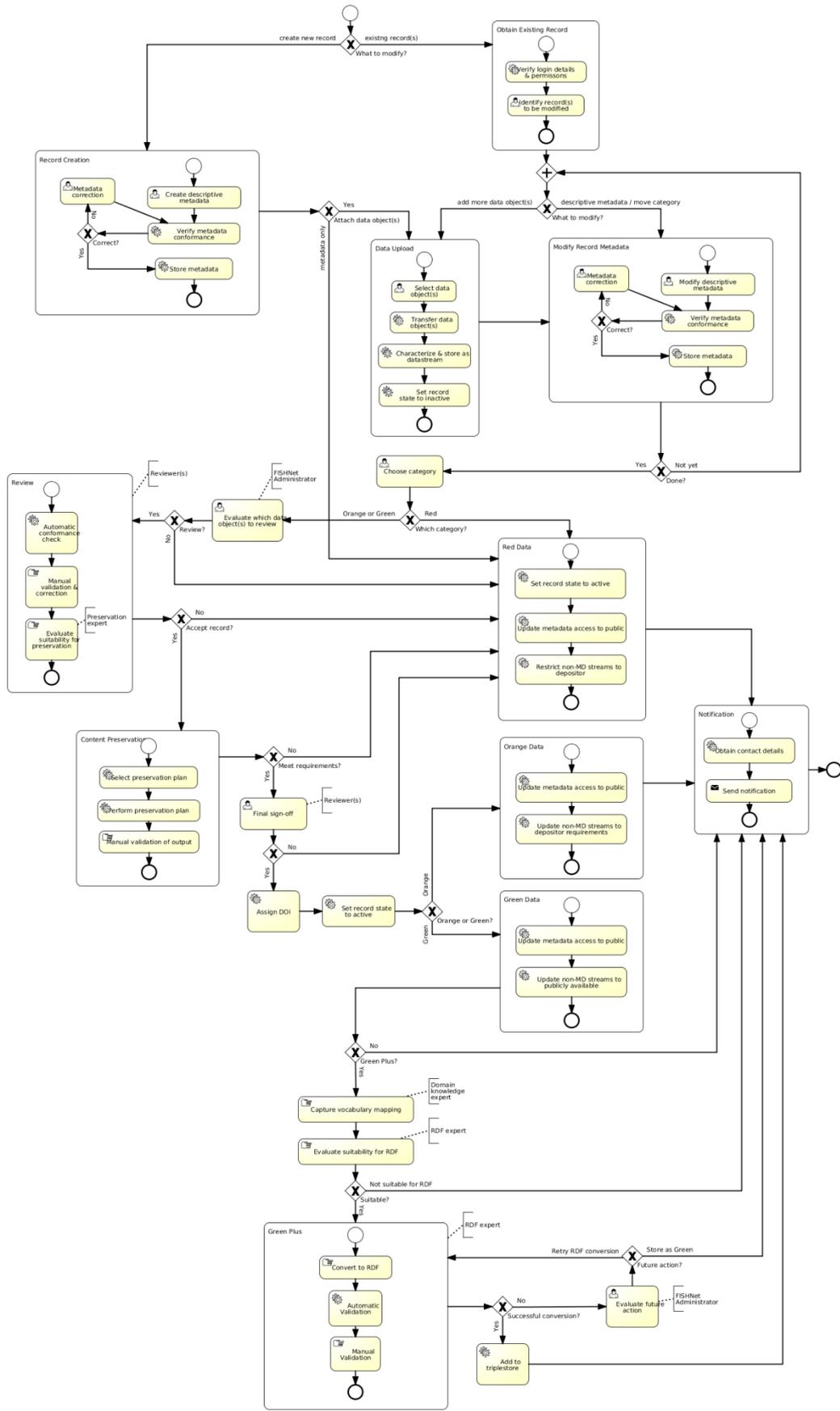


Figure 1: FISHNet Ingest Workflow

Outcomes and Benefits

We envisage that the project will have longer-term impact on the freshwater science community in a number of areas, which we will monitor and evaluate using a variety of metrics.

Improved data capture/creation practices

A key issue affecting the sharing and reuse of data was the failure to apply appropriate quality standards and good practice when creating datasets. In part this was due to an ignorance about such standards and practice within the community, and in part because datasets were created primarily for personal research use, not with a view to reuse and curation, so the application of standards was viewed as an additional burden. It is increasingly accepted that digital curation issues can be addressed most cost-effectively in the early stages of the data lifecycle (Rumsey, 2010).

FISHNet will not support directly the initial data capture or the process of creating the data files; for the current project, this was viewed as “out of scope”, as the FISHNet system’s involvement begins with the deposit of the dataset. Rather, the project’s impact will be, firstly, to educate the community as regards best practice for dataset creation, including the use of appropriate metadata and vocabulary standards, and, secondly, to enforce – or rather encourage – the application of such standards as part of the deposit process, for example by implementing quality assurance for data and metadata as part of the ingest workflow. Repository staff will work with the creators to ensure that datasets conform to appropriate standards and will be readily understandable by other users. As this is likely to require effort on the part of the depositor, incentives will be offered by (i) guaranteeing that appropriate data curation and preservation practices are applied to the data, and (ii) providing DOIs for datasets, facilitating citation and increasing visibility of the depositor’s work.

One outcome of this will be to change the practices of individual researchers when they recording data, by getting them to take into account the requirements for data reuse and curation in addition to their immediate research needs. It is important to note that this is not just a matter of forcing standards on people. While education about generic best practice may act top-down, many of the standards are discipline-specific – for example, formal vocabularies to describe particular domains – and emerge bottom-up from within the community itself. In such cases the repository acts rather as a focus for disseminating good practice, and thus helps the both good practice and the system itself evolve in response to community needs. As datasets submitted to the FISHNet system undergo a quality review as part of the ingest workflow, we can measure our impact in this area by tracking the level of conformance to standards, and conversely the level of changes requested of the depositors, over time.

Data discovery and reuse

Making datasets available through the repository will greatly improve discovery and reuse of data by other potential users in the research and academic communities. Until now sharing in the freshwater community has largely occurred on an ad hoc basis; researcher A hears that researcher B had a particular dataset and asks for a copy. FISHNet provides a discipline-centric and community-centric focus for discovering and accessing such data. A “value-added” aspect of the project is that, as well as researcher-submitted datasets, the repository will integrate FBA-owned datasets, information from the FBA’s extensive freshwater library catalogue, as well as image archives and other sources of freshwater information, encouraging users to regard the repository as a first port of call when seeking freshwater information. We will monitor this type of impact by tracking the levels of access to datasets held in the repository and citation of these datasets by other researchers.

Credit (academic or commercial) for reuse of data

If a dataset submitted to the repository satisfies certain quality and access criteria, we will assign a DOI to it, enabling it to be cited easily in research publications that make use of it. Moreover,

datasets will be made available under straightforward licence agreements that make it clear exactly what the obligations on reuse are, e.g. citation of the original dataset. Again, a potential metric for this type of impact would be an increase in the levels of access to datasets held in the repository and citation of these datasets by other researchers.

Linking data and journals

Discussions have taken place with Dryad⁹, and with the JISC-funded Dryad UK¹⁰ project, concerning mechanisms for linking up peer-reviewed publications in journals and conference proceedings with the data that underlies them (see also Greenberg (2009)). There is potential here for working with journals in the disciplines covered by FISHNet, along the model developed by Dryad, and for integrating publication workflows with the FISHNet data ingest workflow described above. Editorial policies on data deposit have potential for increasing the population of the repository and might also be a route to sustainability. In the first instance, we expect to trial such approaches with the FBA's own portfolio of journals¹¹, and they will also be addressed as part of the subsequent DTC Archive project (see Conclusions and Future Work), and the outcomes fed back into FISHNet.

Longer-term curation and preservation of datasets

Although many UK higher education institutions possess an institutional repository (IR) of some form, these often focus on scholarly works rather than datasets, and in any case their staff are unlikely to possess the expertise necessary to curate datasets in a particular discipline such as freshwater biology at more than a very basic level. Datasets deposited in IRs thus might not be subject to adequate data curation practices, and it would be difficult for researchers to find such datasets as the IRs would be unlikely to support the domain-specific metadata schemas required. On the other hand, while there are subject-based repositories that manage data of interest to freshwater scientists, they do not aim to curate datasets created by individual researchers, who make up a major part of the community. FISHNet has addressed this situation by providing an environment that can provide long-term curation and preservation services to depositors, and facilitate the application of good digital curation practice, thus increasing the longevity of the data. We will track this by monitoring the continued accessibility of the datasets in the long term.

Access control, rights and IPR

There was considerable confusion in the community concerning rights and IPR for datasets, and the project's objective in this area was to encourage researchers to make datasets as freely available as possible, while respecting the researchers' rights to impose restrictions when they deemed it appropriate. The use of the simplified "Traffic Light" system for deposit described above will clarify and simplify licensing of freshwater datasets, and its impact will be measured by monitoring rates of deposit and reuse of datasets, as well as through feedback from data creators and data users.

Sustainability

The FISHNet repository will be managed in the long-term by the Freshwater Biological Association's Data and Information Services staff. The provision of freshwater data and information has been a key part of the FBA's remit since it was founded in 1929, as one of its charitable objectives was to maintain an information store on the subject of freshwater science. The provision of digital data and information is now a key part of the FBA's long-term business strategy, and as a charity, the FBA's position as an independent broker in freshwater matters is one of its unique selling points to the freshwater community. Moreover, the FBA is a membership-based organisation, and thus it is uniquely placed to provide a community-centric focus for data curation and sharing among freshwater scientists. As part of the project, we developed a business case that is aligned with and

⁹ <http://datadryad.org/>

¹⁰ <http://datadryad.org/dryaduk>

¹¹ <https://www.fba.org.uk/journals/>

builds upon the FBA's broad remit and current revenue-generating strategy, as well as its essentially community-focused nature.

Conclusions and Future Work

Some of the conclusions are necessarily tentative until the system has been evaluated adequately in a production environment; however, subject to this proviso, we draw a number of conclusions:

- The project demonstrates the benefits of a simple (from the user's point of view) data archive for disciplines that lack such facilities, in particular the "small" sciences where datasets tend to be hand-crafted by individual researchers or small groups. User engagement carried out by King's with other research groups of this type has revealed similar issues and requirements, leading us to conclude that a generic (not discipline-specific) solution may be possible and desirable, although this may be combined with more discipline-focused components (e.g. to deal with specific metadata standards).
- Assigning DOIs to datasets can provide a motivation to researchers to submit their data to a repository such as FISHNet, by allowing their datasets to be cited more easily and thus raising their academic profile.
- A simplified approach to rights and restrictions, such as the "traffic light" system used by FISHNet, can encourage deposit by researchers who are confused and discouraged by the complexities that often exist around these issues

The boundaries of the FISHNet repository's involvement in freshwater scientists' research processes may be thought of in the following terms:

- Involvement begins when an already-created dataset is deposited with the system. We do not address the creation of the dataset, other than through informal advice and feedback to the depositor.
- Involvement ends when a user has found and downloaded the datasets they require. We do not address what a user does within these datasets, which are regarded by the repository as "black boxes", described only at the dataset level.

This indicates two dimensions in which the scope of the system could be extended:

- Direct support for the capture or creation of datasets, for example via templates or vocabulary services.
- Increased support querying or otherwise working with the datasets, both within individual datasets and across multiple datasets.

The experiences of the FISHNet project have led directly to two subsequent projects that are addressing these issues. FISH.Link¹², which investigated querying across datasets managed by FISHNet using a linked data model, and the DTC Archive project¹³, funded by DEFRA (Department for Environment, Food and Rural Affairs)¹⁴, which is building upon the architectural approach of FISHNet and FISH.Link, but with different data and stakeholders, and more extensive user functionality.

Acknowledgements

The FISHNet project was a collaboration between King's College London, UK and the Freshwater Biological Association, UK, and was funded by the Joint Information Systems Committee (JISC), UK, as part of its Managing Research Data programme.

¹² <http://www.fishlinkonline.org/>

¹³ <http://new.freshwaterlife.org/web/dtc-archive/project>

¹⁴ <http://www.defra.gov.uk/>

References

- Borgman, C. L. (2007), *Scholarship in the digital age: Information, infrastructure, and the Internet* (Cambridge, MA: MIT Press).
- Borgman, C.L., Wallis, J.C., Enyedy, N. (2007), "Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries", *International Journal on Digital Libraries*, Vol. 7, No. 1, 17–30, doi:10.1007/s00799-007-0022-9.
- Borgman, C. L. (2010), "Research Data: Who will share what, with whom, when, and why?", *China-North America Library Conference*, Beijing, <http://works.bepress.com/borgman/238>
- Brase, J. (2009), "Datacite - a global registration agency for research data", *COINFO '09: Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology*, 257-261, <http://dx.doi.org/10.1109/COINFO.2009.66>
- Greenberg, J. (2009), "Theoretical Considerations of Lifecycle Modeling: An Analysis of the Dryad Repository Demonstrating Automatic Metadata Propagation, Inheritance, and Value System Adoption", *Cataloging & Classification Quarterly*, Vol. 47, No. 3, 380-402, DOI: 10.1080/01639370902737547, <http://www.tandfonline.com/doi/abs/10.1080/01639370902737547>.
- Key Perspectives Ltd (2010), "Data dimensions: disciplinary differences in research data sharing, reuse and long term viability. SCARP Synthesis Study", *Digital Curation Centre*, <http://hdl.handle.net/1842/3364>
- Lyon, E., Rusbridge, C., Neilson, C., Whyte, A. (2010), *Disciplinary Approaches to Sharing, Curation, Reuse and Preservation*, *DCC SCARP Final Report*, <http://www.dcc.ac.uk/sites/default/files/documents/scarp/SCARP-FinalReport-Final-SENT.pdf>.,
- Rumsey, A. S. (ed.) (2010), *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*, *Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access*, http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf
- Shearer, K. (2009), *Survey of Digital Preservation Practices in Canada*, *Library and Archives Canada*, <http://www.collectionscanada.gc.ca/digital-initiatives/012018-3100-e.html>